THE INFLUENCE OF RELATIVE INTENSITY ON THE PERCEPTION OF ONSET ASYNCHRONIES

Werner Goebl[1), 2)] *& Richard Parncutt*[1)]

[1)]Department of Musicology, University of Graz
Mozartgasse 3, A-8010 Graz, Austria

[2)]Austrian Research Institute for Artificial Intelligence
Schottengasse 3, A-1010 Vienna, Austria
wernerg@oefai.at; parncutt@kfunigraz.ac.at

ABSTRACT

We address the perception of small onset asynchronies as typically found in expressive piano performance (*melody lead*), which is associated with differences in hammer velocity and hence loudness or salience of chord tones. In three experiments, 26 musicians heard harmonic major-sixth dyads with both tones in the range B4 to Bb5. The tones in each dyad were either both pure, both sawtooth, or both recorded acoustic piano; and either synchronous or asynchronous. First, participants adjusted the relative level of the two tones until they sounded equally loud. This resulted in roughly equal SPL for pure and piano tones, but in the sawtooth tones the higher tone was typically 6 dB more intense, possibly due to simultaneous masking among the partials. In the next experiment, the relative timing and loudness of the two tones were simultaneously manipulated by up to ±54 ms and ±20 MIDI units. The relative perceptual salience of the tones was found to depend on their relative intensity, but not on their asynchrony. Then, in a further experiment, listeners were asked whether the tones were simultaneous, asynchrony was harder to detect when the louder tone began earlier (melody lead). Two possible explanations: either musicians perceive familiar combinations of asynchrony and intensity difference as more synchronous than unfamiliar combinations, or sensitivity to synchrony is reduced in the melody-lead condition by forward masking.

1. INTRODUCTION

A pianist can “bring out” a melody tone (increase its perceptual *salience*) either by depressing the key more quickly or by varying timing relative to the accompaniment. Melody tones typically sound some 30 ms before the other tones of a chord (*melody lead*, Palmer, 1989, 1996; Repp, 1996; Goebl, 2001); this effect is generally associated with, and presumably causally related to, differences in hammer velocity between the melody and accompaniment (*velocity artifact*, Repp, 1996; Goebl, 2001). Here, we investigate some perceptual aspects of this phenomenon.

In preliminary experiments on the perception of harmonic dyads (Goebl & Parncutt, 2001), we found no significant difference between the perceptual prominence of a delayed and an anticipated higher tone. This result cast doubt on the frequently encountered (and often tacit) assumption in the piano performance literature that melody lead attracts more attention to the melody than melody lag. We also found that participants could report the correct order of two (equally intense) stimuli at asynchronies exceeding about ±40 ms, irrespective of tone complexity.

In the present study, we first asked which relative level of the tones of a harmonic major-sixth dyad produces an impression of equal loudness or salience. This was done separately for each listener and for each of three different tone types. Using these data as a baseline, we then investigated the relative perceptual salience of the tones of a harmonic dyad in which timing and intensity were simultaneously varied. Finally, we investigated listeners' sensitivity to asynchrony in the context of changes in relative intensity.

2. EXPERIMENT I

2.1. Method

Participants. The 26 participants were aged between 23 and 32 years. All were musicians who had been playing their instrument regularly for an average of 18.9 years. 23 of them had studied their instrument at post-secondary level for an average period of 8.3 years. They comprised 15 pianists, 5 violinists, 1 singer (a tenor), 3 cellists, 1 double bass player, and 1 composer (who regarded the computer as his main instrument).

Stimuli. Three tone types were used: pure, sawtooth with 16 partials (-6 dB per octave), and piano. To avoid uncontrolled asynchronies and sympathetic vibrations, harmonic dyads of piano tones were created by digital superposition of individual monophonic tone recordings. The MIDI velocity values ranged from 79/31 (higher/lower tone) to 31/79, in increments of ±2 units (79/31, 77/33, 75/35, etc.). The nominal equality was thus 55/55 — a typical *mezzoforte*. The amplitudes of the pure and sawtooth stimuli were similar to those of the recorded piano sounds.[1]

Five different dyads were presented to the participants, each spanning the musical interval of a major sixth. Three of them comprised piano tones: B4 and G#5 (approx. 494 and 831 Hz), C5 and A5 (523 and 880 Hz), and Db5 and Bb5 (554 and 932 Hz) respectively.[2] The other two dyads were synthetic; one comprised two pure tones, the other two sawtooth tones. In both cases the (fundamental) frequencies were 523 and 880 Hz, corresponding to C5 and A5.

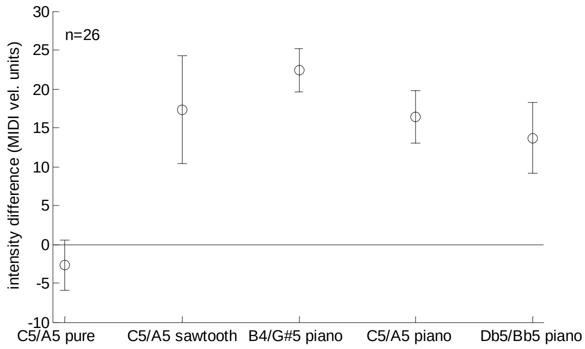
Equipment. The acoustic piano tones were played on a computer-controlled Bösendorfer SE290 (at every MIDI velocity between 20 and 90[3]) and recorded with AKG (CK91) micro-

[1] The relationship between key velocity (in MIDI velocity units) and peak SPL (in dB) for 1700 single tones played on the Bösendorfer SE290 was approximated by the expression: $-77.2 + 26.1 \cdot \log_{10}(\text{MIDI velocity}) + 5.3 \cdot \log_{10}(\text{MIDI velocity})^2$.

[2] The given frequency values are calculated and correspond to equal temperament with a A4 at 440 Hz. The actual frequency of the lowest partial of each tone will be slightly different from these values due to inharmonicity and pitch shifts.

[3] In this study, the relation between MIDI velocity and hammer velocity (in meters per second) at the Bösendorfer system was set to be: MIDI velocity = $52 + 25 \cdot \log_2(\text{hammer velocity})$.

a)



b)

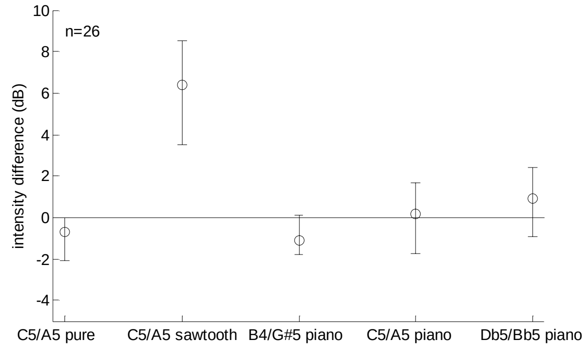


Figure 1: Intensity difference between simultaneous tones judged to be equally loud, averaged across 26 participants. Bars: 95% confidence intervals. Vertical axes: **a)** relative to MIDI velocity or equivalent; **b)** in dB. In each case, a positive value means that the higher tone was more intense than the lower at equal loudness.

phones (placed approx. 1 m from the strings, in a 6 by 6 meter room) onto a Tascam DA-P1 DAT recorder. They were transferred digitally to the hard disc of a PC using a Creative SB live! 5.1 Digital soundcard and stored in WAV format (16-bit, 44.1 kHz). The pure and synthetic complex tones were generated using Matlab® software. During the experiment, all sounds were played back via the same soundcard and Sennheiser HD 25-1 headphones (diotic presentation: same signal in each ear). The experiments were controlled by a computer program that had been developed by the first author in a Matlab® environment.

Procedure. In each trial, participants adjusted the level of two simultaneous tones relative to each other until they sounded equally loud. Five trials were presented in a random order that differed from one participant to the next. The relative intensities of the two tones at the start of each trial was also selected at random, from 25 possibilities. Participants first adjusted the relative level of the tones in relatively large increments of ±6 MIDI units (i.e., one tone became 6 units louder while the other became 6 units softer, so that the difference in MIDI velocities changed by 12 in each step). In the second block, the five stimuli were repeated in the same order and adjusted in smaller steps of ±2 MIDI units. Participants were asked to go past the point of equal loudness and return to it from the other side before going on to the next dyad. Each stimulus could be adjusted and repeated for an indefinite period. To test reliability, the entire procedure (coarse followed by fine tuning) was repeated. If the mean difference between the results for the first and second block was larger than 6 MIDI velocity units, a third block was run; this happened for 5 of the 26 participants. After all three experiments were completed, a questionnaire was filled in. The whole session lasted between 30 and 70 minutes. Participants were paid 20 Euro for their services.

2.2. Results & Discussion

The mean adjustments over all 26 participants' median adjustments are plotted in Figure 1a against the (equivalent) MIDI velocity differences between the two tones (see footnote 1). A positive difference on the *y* axis indicates that, at equal loudness, the higher tone had greater intensity or MIDI velocity. The data initially suggest that for all three piano dyads and for the sawtooth dyads, the higher tones had considerably higher level than the lower tones at equal loudness (salience). Piano tones with constant hammer velocities do not necessarily have the same peak SPL.[4] For example, B4 on our piano samples was always 6 dB more intense than G#5 played with the same MIDI velocity.[5] Once the data have been adjusted to account for this (Figure 1b), the sound level differences in the piano samples disappeared. Only in the sawtooth sounds was there a significant difference in SPL (of about 6 dB) at equal subjective loudness.

The effect cannot be accounted for by the Fletcher-Munson loudness curves for pure tones, which would predict just the opposite tendency. Instead, the effect may be accounted for by masking between the higher partials. Since lower pure tones generally mask higher pure tones more than the reverse (Moore, 1997), higher harmonic complex tones may need to have greater SPL to be perceived as equally loud as simultaneously sounding lower complex tones with identical temporal and spectral envelopes. The masking effect in the piano tones may have been less prominent due to the spectral and temporal variability of the amplitudes of the partials, and because the spectral slope of each tone depends on both intensity and register. The greater spread in the data for the sawtooth tones is consistent with comments on the final questionnaire to the effect that the sawtooth sounds were the hardest of the three tone types to judge, presumably due to unfamiliarity.

3. EXPERIMENT II

3.1. Method

Equipment and the participants were the same as in the previous experiment. Each of the three tone types (pure, sawtooth, and piano) was presented in five intensity combinations and with five degrees of asynchrony, resulting in $3 \times 5 \times 5 = 75$ stimuli. The intensity combinations were +20/-20, +10/-10, 0/0, -10/+10, and -20/+20 MIDI velocity units, relative to the median levels judged to be equally loud in the previous experiment; the baselines were maintained separately for each tone type and for each participant. The asynchronies were -54, -27, 0, 27, and 54 ms (where a negative value indicates that the higher tone began be-

[4] The peak dB value of a piano sample was the maximum value of a RMS smoothed sound signal. The window size was 10 ms.

[5] The peak dB values for the same key and the same hammer velocity change strongly with microphone position. The iso-velocity lines of the second channel of our piano recordings showed a quite different picture than the first channel (for a similar discussion see Repp, 1997, p. 1880). The entire phenomenon will be discussed in detail elsewhere.

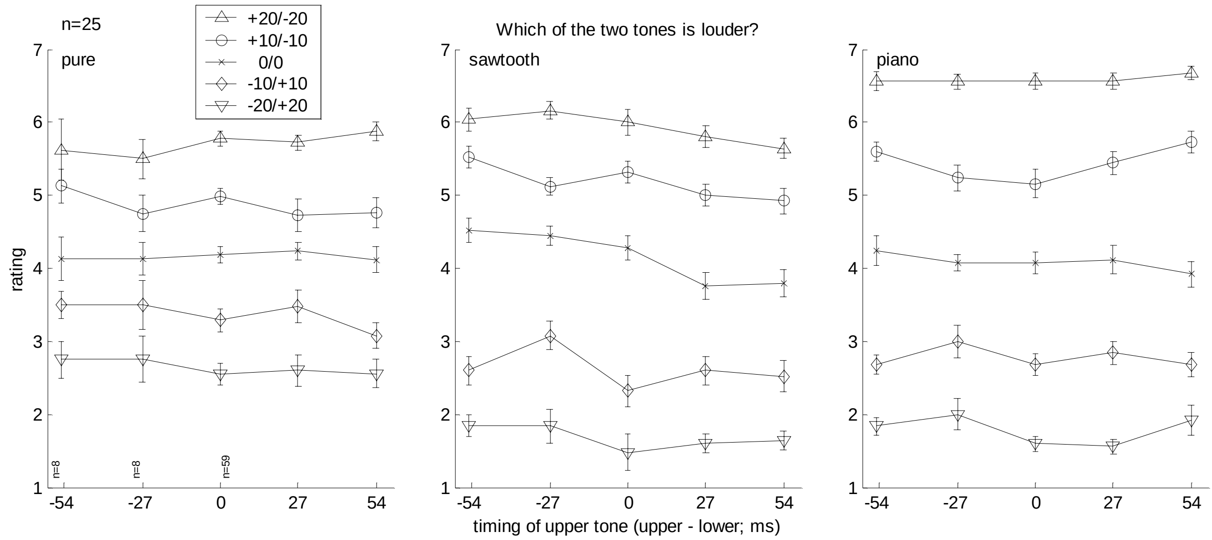


Figure 2: Mean relative loudness ratings of 25 participants in Experiment II, as a function of tone type and asynchrony. Rating scale: 1, upper tone much louder; 4, two equally loud; 7, lower tone much louder. The five horizontal lines in each panel correspond to the 5 intensity combinations (MIDI velocity of upper tone relative to lower tone); the three panels correspond to the three tone types (pure, sawtooth, piano). Error bars are 95% confidence intervals of the means.

fore the lower tone). Regardless of whether the onset was synchronous or asynchronous, the tones always sounded together for a total of 350 ms, and faded out simultaneously.

The chosen time differences were typical of melody leads in piano performance. The velocity artifact hypothesis (Repp, 1996; Goebl, 2001) is based on the simple observation that the faster a piano key is depressed, the earlier the hammer arrives at the strings. In a typical modern grand piano, when two tones are struck simultaneously from the key surface, a higher tone that is 20 MIDI velocity units louder than the lower tone typically sounds about 27 ms before the lower tone.

Participants indicated which of the two tones sounded louder on a scale from 1 (lower tone much louder) to 7 (higher tone much louder). Equal loudness was indicated by 4. After 13 practice stimuli, the 75 stimuli were presented in a random order that was varied from one listener to the next. Each stimulus could be repeated as often as desired before deciding on a rating.

3.2. Results & Discussion

In the final questionnaire, participants indicated that this experiment was the most difficult of the three. One participant's results had to be excluded, because he could not perform the task at all (as he indicated in the questionnaire).

The mean ratings are plotted in Figure 2.[6] A 3-way ANOVA on the ratings with timing, intensities, and timbre as factors revealed significant main effects of relative intensity and timbre (tone type), but not of timing (synchrony). Regarding timbre, the range of judgements was smallest for pure tones and largest for the piano tones, suggesting that the difference in salience between two simultaneous tones depends on the number of audible harmonics in each tone (probably consistent with the masking hypothesis advanced above). Surprisingly, there was no main effect of timing: the lines in Figure 2 do not depart significantly from horizontal.

4. EXPERIMENT III

The psychoacoustic literature initially suggests that listeners can easily distinguish synchronous from asynchronous dyads: the temporal order threshold (which tone came first?) is around 20 ms (Hirsh, 1959), and the threshold for asynchrony detection (were the tones synchronous?) can be as low as 2 ms (Zera & Green, 1993). But in (piano) music, where tones have overlapping spectra and unequal loudness (so that one tone is masked by the other), the thresholds are higher. In this experiment we set out to measure these higher thresholds, while simultaneously varying the relative sound level of the tones.

4.1. Method

The procedure, participants, and stimuli were identical to Experiment II. The only difference was the question: The participants were asked whether or not the two tones were simultaneous, in a 2AFC paradigm.

4.2. Results & Discussion

The results are plotted in Figure 3. The expected responses were 'yes' (1) for synchrony, and 'no' (0) for the four asynchronous conditions. The synchronous dyads were reliably recognized for the two artificial tone types, independent of relative intensity. But for the piano tones, asynchronous dyads were often heard to be synchronous when the louder tone preceded the softer tone (melody lead). This effect also appeared for the synthetic tones; it was weakest for the pure tones and strongest for the piano sounds. For instance, the +20/-20 condition (upwards triangle in Figure 3) at –27 ms (sawtooth and piano) was perceived as simultaneous by around 70% of the participants, whereas at +27 ms it was heard as asynchronous by almost everyone.

[6] Due to a programming mistake of the first author, the -54 and the -27 ms condition of the pure tones were omitted and the 0 ms condition presented three times for the first 17 participants (also in Exp. II). The modified n values are specially indicated in Figures 2 & 3, the different lengths of the error bars reflect this.

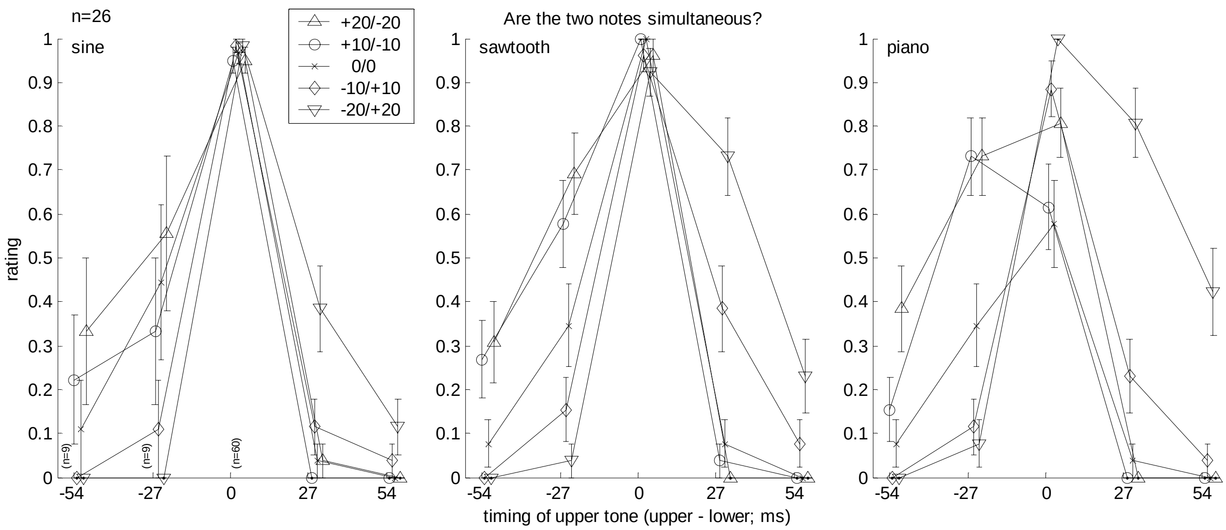


Figure 3: Experiment III. Mean ratings over 26 participants as a function of asynchrony. The answers could be yes (1) or no (0). Separate lines for the different intensity combinations (relative MIDI velocity upper/lower tone), separate panels for the three tone types. The errorbars indicate standard error at 95% confidence level. (The symbols with the error bars are shifted horizontally within each category for the sake of visual clarity.)

Two possible explanations may be advanced for these asymmetries. The first involves familiarity with piano music: either listeners are insensitive to melody lead in piano performance (due to overexposure); or the listeners in our experiment only noticed, and hence correctly identified, asynchrony when a musically unfamiliar combination of relative loudness and timing was presented. Those participants who were also pianists might additionally have associated the characteristic sound of melody lead with the (kinesthetic) sensation of fingers simultaneously striking the key surface. The second explanation is more psychoacoustic in nature: the effect could be due to forward masking. A louder, anticipated tone masks a softer tone by forward masking, which is stronger than backward masking and attenuates the following softer tone for about the same period of time as typical melody leads show (some tens of milliseconds, Zwicker & Fastl, 2001). Simultaneous masking applies especially among the partials of complex tones spanning typical music intervals, consistent with the finding that the observed asymmetry is stronger for complex than for pure tones.

5. CONCLUSIONS

These new insights into the perception of asynchronous onsets with variations in loudness seem to explain why pianists are largely unaware of the melody lead (Parncutt & Holming, 2000). Our experiments need to be repeated with sonorities of more than two tones to check whether, in that case, asynchrony affects salience, as predicted by Bregman's (1990) theory of auditory scene analysis, and whether melody lead and melody lag have unequal effects on melodic salience.

6. ACKNOWLEDGEMENTS

This research is part of the project Y99-INF, sponsored by the Austrian Federal Ministry of Education, Science and Culture (BMBWK) in the form of a START research Prize to Gerhard Widmer. The authors thank the Acoustics Research Institute of the Austrian Academy of Sciences for generously providing technical equipment and the Bösendorfer company for providing the SE290 grand piano for experimental use. Thanks to Bernhard Laback and Simon Dixon for helpful comments.

7. REFERENCES

Bregman, A. S. (1990). *Auditory scene analysis. The perceptual organization of sound*. Cambridge, Massachusetts: The MIT Press.

Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America, 110*(1), 563-572.

Goebl, W., & Parncutt, R. (2001). Perception of onset asynchronies: Acoustic piano versus synthesized complex versus pure tones. *Meeting of the Society for Music Perception and Cognition (SMPC2001), 9-11 August 2001*, 21-22.

Hirsh, I. J. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America, 31*, 759-767.

Moore, B. C. J. (1997). *An Introduction to the psychology of hearing* (4th ed.). San Diego, California: Academic Press.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 331-346.

Palmer, C. (1996). On the assignment of structure in music performance. *Music Perception, 14*, 23-56.

Parncutt, R., & Holming, P. (2000). Is scientific research on piano performance useful for pianists? *Poster presentation at the 6th international conference on music perception and cognition (ICMPC'2000)*, Keele, UK.

Repp, B. H. (1996). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America, 100*(6), 3917-3932.

Repp, B. H. (1997). Acoustics, perception, and production of legato articulation on a computer-controlled grand piano. *Journal of the Acoustical Society of America, 102*(3), 1878-1890.

Zera, J., & Green, D. M. (1993). Detecting temporal onset and offset asynchrony in multicomponent complexes. *Journal of the Acoustical Society of America, 93*(2), 1038-1052.

Zwicker, E., & Fastl, H. (2001). *Psychoacoustics. Facts and Models*. Berlin, Heidelberg: Springer.